DOCUMENT RESUME

ED 117 118 TM 004 574

AUTHOR Olson, George H.
TITLE Applications of the Multivariate General Linear Hypothesis in Educational Research and Evaluation.
PUB DATE 31 Mar 75
NOTE 37p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS *Analysis of Covariance; *Analysis of Variance; Data Analysis; *Educational Research; *Hypothesis Testing; Mathematical Applications; Mathematical Models; Matrices; *Statistical Analysis
IDENTIFIERS *Multivariate General Linear Hypothesis

ABSTRACT

The multivariate general linear hypothesis (MGLH) has received relatively little utilization in educational research and evaluation. This is surprising in view of the fact that recent publications have made the MGLH tractable by practitioners. This paper seeks to stimulate interest in the MGLH by reviewing recent applications, emphasizing the advantages of the MGLH as a general data analytic tool, and making suggestions for conducting research within an MGLH framework. (Author)

Applications of the Multivariate General Linear Hypothesis in Educational Research and Evaluation[1]

George H. Olson
Dallas Independent School District

By now, researchers in the educational milieu should be disposed to the idea that both causes and effects of educational phenomena are inherently multivariate in nature (see, for example, Tatsuoka's recent review [1973] as well as older surveys of multivariate statistical applications in education and psychology contained in books by Cattell [1962] and Whitla [1968]). In their statistical applications, however, educational researchers have apparently taken only half of this statement seriously. Thus, researchers have recognized the fact that causes of educational phenomena are multidimensional and have employed techniques of multiple linear regression with increasing levels of sophistication to investigate both the combined effects of, and interrelationships among, multiple independent variables. For the most part, however, researchers have not investigated the "multivariateness" of educational outcomes. Instead, the typical approach has been to study the effects of a common set of independent variables on each of several criterion variables separately.

This practice is not unique to education. Lana and Lubin (1963), for instance, reviewed the published articles in over three years worth of three APA journals in an effort to discover the frequency with which multiple-criteria designs were used and how they were analyzed. One of their findings, of interest here, was that about one-third of the studies used designs which involved multiple, correlated criteria; yet only one of the studies reviewed took the

[1] Paper presented at the annual meeting of the American Educational Research Association, Session 3.09, Washington D.C., March 31, 1975.

correlation among the criterion variables into account. Multiple criteria were usually analyzed by applying analysis of variance techniques to each variable separately.

In preparing for this paper, I did a quick, informal review of the studies published in the last three years in the <u>American Educational Research Journal</u> and the <u>Journal of Educational Psychology</u>. Over 80 percent of the experimental and comparative studies reported in these journals investigated effects on multiple criteria. Although several of these studies employed multivariate techniques, the majority of them (approximately three-fourths) employed multiple univariate analyses of variance.

The typical example would be a study in which several independent variables (including pretests, demographic data, and incidence of treatment measures) are examined separately for their effect on, say, measures of motivation, achievement, and attitude toward school--variables that few would deny are correlated (perhaps highly) in the population. The possible ramifications of performing separate univariate analyses on correlated criteria are well-known (e.g., Hummel and Sligo [1971]). If a Type I error occurs in tests involving one criteria, the probability is greater than α that it will occur in the tests involving the other criteria also. This fact, by itself, should provide sufficient motivation to the researcher to seek multivariate techniques for the analysis of multiple-criteria designs. At the very least a multivariate technique offers the researcher a procedure for controlling experiment-wise probability of a Type I error.

Of course, there are other reasons for calling for an increased application of multivariate methods. For instance, Snow (1974) has recently argued that research designs in education need to be more representative of possible outcomes. That is, researchers need to choose samples of dependent variables which are representative of the phenomena being studied. Snow argues effectively that if the results of educational experiments are to be generalizable, they must first be representative. Part of Snow's thesis can be interpreted as a call for multivariate investigations in educational research. Multivariate statistical procedures appropriately fit this epistemological point of view that the "effects" in educational settings are rarely, if ever, unidimensional and therefore should not be studied in isolation. The study of multiple outcomes *simultaneously* can afford the researcher an opportunity to "uncover" complex relationships among treatment and outcome variables that might otherwise go undiscovered. An example might be a situation in which it is found that two experimental instructional programs lead not only to increases in several measures of achievement and motivation, but also, under one of the programs, to increases in the correlation among the variables. Such a finding might have important implications for understanding the dynamics of the instructional programs.

If the need for greater application of multivariate statistical procedures can be established so easily, why then are these procedures not used more frequently in educational research? Although it is impossible to answer this question completely, there is no doubt that at least part of the answer lies in the fact that multivariate

procedures are more complex, both mathematically and conceptually, then univariate procedures. Until recently, researchers who sought to use multivariate techniques had to first develop a fairly high level of mathematical skill to be able to read the existing textbooks. Even given the requisite mathematical ability, the existing textbooks were often less than useful to the applied researcher since they contained a paucity of examples on anything even bearing a passing resemblance to the kinds of problems encountered in educational research.

In recent years, however, a number of intermediate-level (in mathematics) books on multivariate statistics have been published (e.g., Cooley & Lohnes [1971], Finn [1974], Harris [1975], Morrison [1967], Press [1972], Tatsuoka [1971], and Van de Geer [1971]). Although these books usually require some ability in matrix algebra, the mathematical developments tend to be highly tractable. More than that, however, these books provide a rich source of examples of applications of multivariate procedures. It is likely that, due to these books alone, many more applications of multivariate procedures will appear in the published educational research literature in the next few years.

Another development of recent years that is likely to lead to a greater use of multivariate procedures is the increase in the number of computer programs that have become available. Many of the books listed above (especially in Harris [1975], Press [1972], and Cooley & Lohnes [1971]) provide references to existing computer programs. The availability of these programs has further lessened the

demand on applied researchers for mathematical sophistication. Of course, the researcher still needs a conceptual understanding of the particular procedures he intends to use. Providing the researcher has a clear understanding of what he wants to do statistically, as well as a conceptual understanding of the statistical techniques he intends to employ, then given a computer program for the statistical procedure, a multivariate analysis is relatively straight-forward. This is especially true for most applications of the multivariate general linear hypothesis (MGLH), the subject of this paper to which we now turn.

The MGLH is the most general of all parametric linear statistical procedures. In fact, all linear statistical tests (univariate and multivariate) can be developed as special subclasses of MGLH theory. This includes, among other procedures, factor analysis, discriminant analysis, prediction (or regression) analysis, and the analysis of variance and covariance. Although the mathematical development of MGLH theory is complex, its conceptualization, at least insofar as the most common applied situations are concerned, is not particularly difficult. It does require an elementary facility for matrix algebra, however. A simple statement of the theory is that if a set of dependent variables, $\underset{\sim}{Y}$ (where the tilde underscore denotes a matrix), is linearly related to a set of independent variables, $\underset{\sim}{X}$, by the equation,

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{E}, \qquad [1.1]$$

where $\underset{\sim}{\beta}$ is a set of unknown parameters of interest, then <u>any</u> hypothesis of the form

$$H_o: \underset{\sim}{A}\underset{\sim}{\beta}\underset{\sim}{C} = \underset{\sim}{D} \qquad [1.2]$$

is testable. In this equation, the matrices $\tilde{A}$ and $\tilde{C}$ are used to select particular elements from among rows and columns of $\tilde{\beta}$. The matrix, $\tilde{D}$, is a matrix of constants specified by the investigator. Usually $\tilde{D}$ is set equal to $\tilde{0}$, a matrix of zeros.

In the next section a brief introduction to the theory of the MGLH is provided. Readers who wish to pursue this development in greater depth are encouraged to consult the appropriate chapters in Kempthorne (1952), Kulbach (1968), Mendenhall (1968), and Seal (1964). Early papers by Smith, Gnanadesikan, and Hughes (1962) and Bock (1963b) develop the MGLH in a way that is useful to those who may wish to program the computations. Thorough, but highly mathematical, presentations of the MGLH have been given by Anderson (1958), Rao (1965), and Seber (1966).

In the third section of this paper, several general algebraic examples of typical applications have been provided. Other examples of applications of the MGLH, using real data, can be found in Bock (1963a), Bock and Haggard (1968), Jones (1966), and Finn (1974).

2. *Introduction to the Theory of the MGLH*

In this paper, interest is primarily focused upon the analysis of variance and the analysis of covariance. Repeated measures (or profile) analyses of variance and covariance are also covered but are treated as special instances of multivariate analyses of variance and covariance. Before proceeding to illustrative applications of the MGLH, however, it is first necessary to lay some of the mathematical ground work.

Definition of the Model

We begin by rewriting the general equation for a multivariate linear model,

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{E} \qquad [2.1]$$

where the component matrices are described in the paragraphs which follow.

$\underset{\sim}{Y}$ is an $N \times p$ matrix consisting of $1 \times p$ response row vectors for each of N subjects. An element in $\underset{\sim}{Y}$, denoted as $Y_{ij}^{(\ell)}$, would constitute the ℓ'th response measure taken on the i'th individual in group j.

$\underset{\sim}{X}$ is an $N \times q$ matrix of predictor and/or design variables. In multiple linear prediction, $\underset{\sim}{X}$ would contain the $1 \times q$ vectors of predictor scores for each of the N subjects. In analysis of variance, $\underset{\sim}{X}$ would be a matrix of design variables. In the analysis of covariance, $\underset{\sim}{X}$ would contain a combination of predictor scores and design variables.

$\underset{\sim}{\beta}$ is a $q \times p$ matrix of unknown parameters, the elements of which are of interest in tests of MGLHs. Depending upon the choice of $\underset{\sim}{X}$,

the elements of $\underset{\sim}{\beta}$ may represent subgroup means (expected values), contrasts among subgroup means, population regression coefficients, etc.

$\underset{\sim}{E}$ is an $N \times p$ matrix whose columns, $\underset{\sim}{E}^{(\ell)}$, are error vectors for the p response measures. Stated differently, for each row vector of response observations, $\underset{\sim}{Y}'_{ij}$, there is a corresponding row vector, $\underset{\sim}{E}'_{ij}$, of disturbances usually due to errors of measurement and lack of linear fit to the model.

Assumptions

The assumptions applicable to the model defined above, for purposes of testing hypotheses, are the following:

1. The model is linear in terms of the parameters, $\underset{\sim}{\beta}$.
2. $N \geq p + q$.
3. $\underset{\sim}{X}$ is of full rank, q.
4. Each row vector, $\underset{\sim}{E}'_{ij}$, in $\underset{\sim}{E}$, is independently sampled and distributed multivatiate normally (MVN) with expected value, $\underset{\sim}{0}'_{ij}$, and covariance matrix $\underset{\sim}{\Sigma}$

The first assumption is rarely limiting in educational research and evaluation. Many models which appear non-linear on first sight are actually *intrinsically* linear. In these situations, suitable linear models can be written following acceptable transformations of the original measures (see Draper and Smith, 1968; ch. 5, for a discussion of these types of models). Assumption

number 2 is required to ensure the availability of a sufficient number of observations to estimate the pq elements in $\underset{\sim}{\beta}$.

The third assumption is necessary to ensure a unique solution for $\underset{\sim}{\beta}$. Since this solution requires the inverse of $\underset{\sim}{X}'\underset{\sim}{X}$, it is necessary that $\underset{\sim}{X}'\underset{\sim}{X}$ be nonsingular; hence $\underset{\sim}{X}$ must be of full rank.

More general, non-unique solutions for $\underset{\sim}{\beta}$ exist. These solutions, which use generalized inverses of $\underset{\sim}{X}'\underset{\sim}{X}$, allow $\underset{\sim}{X}$ to have more columns than its rank (i.e., they allow $\underset{\sim}{X}$ to be less than full rank). Since full-rank design matrices are usually quite easily constructed for most designs in educational research, the generalized inverse approach is rarely, if ever, needed. Anyway, for any design matrix less than full rank, there always exists a transformation matrix, $\underset{\sim}{T}$, such that

$$\underset{\sim}{X}^* = \underset{\sim}{X}\underset{\sim}{T}$$

is of full rank (see Bock [1963], Graybill [1961; pp. 235-239], or Smith [1972], for details on computing $\underset{\sim}{T}$).

The fourth assumption provides the foundation for the theory underlying tests of MGLHs. The assumption implies that the

$$\underset{\sim}{Y}'_{ij} \sim MVN(\underset{\sim}{\mu}'_{ij}, \underset{\sim}{\Sigma}).$$

The fourth assumption may be stated in an alternative form which will prove useful in the development. Thus, if the columns in $\underset{\sim}{E}$ (and, correspondingly, the columns in $\underset{\sim}{Y}$) are strung out to form the $Np \times 1$ vector $\underset{\sim}{V}_e$ (and $\underset{\sim}{V}_y$), then assumption number 4 states that

$$\underset{\sim}{V}_e \sim MVN(\underset{\sim}{0}, \underset{\sim}{\Sigma} \times \underset{\sim}{I})$$

$$\underset{\sim}{V}_y \sim MVN(\underset{\sim}{V}_\mu, \underset{\sim}{\Sigma} \times \underset{\sim}{I}) \qquad [2.2]$$

where $\underset{\sim}{0}$ is an $Np \times 1$ vector of zeros, $\underset{\sim}{V}_\mu$ is an $Np \times 1$ vector of expected values, $\mathcal{E}\, \underset{\sim}{V}_y$, $\underset{\sim}{I}$ is an $N \times N$ identity matrix, and the symbol, $\times$, denotes the Kronecker direct product (Cornish [1957], Searle [1966, pp. 215-220]; Vartak [1955]). Since the parameters in $\underset{\sim}{\beta}$ are linear functions of $\underset{\sim}{Y}$, the variance of $\underset{\sim}{\beta}$ and, thus, tests of hypothesis involving $\underset{\sim}{\beta}$ follow more or less directly.

Estimation of $\underset{\sim}{\beta}$ *and* $\underset{\sim}{\Sigma}$

Since the elements of $\underset{\sim}{\beta}$ and $\underset{\sim}{\Sigma}$ are not generally known in practice, they must be estimated. Two procedures are available, viz. maximum likelihood (ML) and Least Squares (LS). Since the former requires the assumption of multivariate normality whereas the latter does not, and since the solutions for $\underset{\sim}{\beta}$, are the same in either case, only the LS procedure will be pursued.

The LS procedure calls for obtaining estimates of $\underset{\sim}{\beta}$, $\underset{\sim}{B}$ say, such that the error sums of squares and crossproducts (SSCP) are a minimum. In matrix terms, estimates of $\underset{\sim}{\beta}$ (viz, $\underset{\sim}{B}$) are obtained such that the following equation is a minimum.

$$\underset{\sim}{E}'\underset{\sim}{E} = (\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta})'(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta}). \qquad [2.3]$$

LS procedure proceeds by differentiating the right-hand side of Equation 2.3 with respect to elements within rows of $\underset{\sim}{\beta}$ and setting the result equal to the null matrix (a comformable matrix of zeros).

This procedure yields a system of qp simultaneous equations in $\underset{\sim}{\beta}$ which are jointly set equal to 0. Appropriate manipulation of these equations leads to the system of Normal Equations,

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{B} = \underset{\sim}{X}'\underset{\sim}{Y} \qquad [2.4]$$

which, assuming $\underset{\sim}{X}'\underset{\sim}{X}$ is nonsingular, is easily solved:

$$\underset{\sim}{B} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}. \qquad [2.5]$$

The estimates, $\underset{\sim}{B}$, have been shown to be unbiased and minimally dispersed by several authors (e.g., Anderson [1958], Kulback [1968], Press [1972], Rao [1965]).

A sample estimate of $\underset{\sim}{\Sigma}$, $\underset{\sim}{S}$ say, is obtained ~~form~~ the error SSCP matrix. Thus,

$$\underset{\sim}{E}'\underset{\sim}{E} = \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{B}'\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{B} \qquad [2.6a)$$

or, since $\underset{\sim}{B} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}$,

$$\underset{\sim}{E}'\underset{\sim}{E} = \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{B}'\underset{\sim}{X}'\underset{\sim}{Y}. \qquad [2.6b]$$

This quantity, which we will call $\underset{\sim}{SS}_E$, is then used as an estimate of $\underset{\sim}{\Sigma}$, i.e.,

$$\underset{\sim}{S} = \text{est}(\underset{\sim}{\Sigma}) = (N-p)^{-1}\underset{\sim}{SS}_E. \qquad [2.7]$$

Expectation and Variance of $\underset{\sim}{B}$

From the solution of the normal equations, we obtained

$$\underset{\sim}{B} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}. \qquad [2.5]$$

We note that $\underset{\sim}{B}$ represents a straight-forward linear transformation of $\underset{\sim}{Y}$ which for the time being may be written as

$$\underset{\sim}{B} = \underset{\sim}{L}'\underset{\sim}{Y}, \text{ where } \underset{\sim}{L}' = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'. \qquad [2.8]$$

Taking the expectation of B, we have

$$\mathcal{E}(\underset{\sim}{B}) = \mathcal{E}(\underset{\sim}{L}'\underset{\sim}{Y}) = \underset{\sim}{L}'\mathcal{E}(\underset{\sim}{Y}).$$

From Equation 2.1, however, we have

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{E}.$$

Under assumption 4, the $\underset{\sim}{E}_{ij}$ in E are distributed multivariately about $\underset{\sim}{0}$. Hence,

$$\mathcal{E}(\underset{\sim}{Y}) = \underset{\sim}{X}\underset{\sim}{\beta}.$$

Therefore,

$$\begin{aligned} \mathcal{E}(\underset{\sim}{B}) &= \underset{\sim}{L}'\underset{\sim}{X}\underset{\sim}{\beta} \\ &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{\beta} \\ &= \beta, \end{aligned}$$

which demonstrates the unbiasedness of $\underset{\sim}{B}$.

To develop the covariance matrix of $\underset{\sim}{B}$ we make use of operations on Kronecker products.

From Equation [2.8] we have

$$\underset{\sim}{B} = \underset{\sim}{L}'\underset{\sim}{Y}.$$

If the columns of $\underset{\sim}{Y}$ are strung out to form the Np × 1 column vector, $\underset{\sim}{V}_y$, as before, then this equation may be written as

$$\underset{\sim}{V}_b = (\underset{\sim}{I} \times\cdot \underset{\sim}{L}')\underset{\sim}{V}_y \qquad [2.9]$$

where $\underset{\sim}{I}$ is p × p and $\underset{\sim}{L}'$ is q × N. The direct product, $\underset{\sim}{I} \times\cdot \underset{\sim}{L}'$ is pq × pN; thus, the overall product in Equation 2.9 is pq × 1. From the rules pertaining to the variance of linear transformations of random variables, we have

$$\mathrm{Var}\,[(I \times\cdot L')V_y]$$

$$= (I \times\cdot L')'\ (\Sigma \times\cdot I)\ (I \times\cdot L')$$

$$= (I' \times\cdot L)\ (\Sigma \times\cdot I)\ (I \times\cdot L')$$

$$= (I'\Sigma \times\cdot LI)\ (I \times\cdot L')$$

$$= I'\Sigma I \times\cdot LIL'$$

$$= \Sigma \times\cdot (X'X)^{-1}.$$

Thus, for the distribution of B, we may write

$$B \sim \mathrm{MVN}[V_\beta;\Sigma \times\cdot (X\ X)^{-1}].$$

Through similar logic it can be shown that for linear transformations of B,

$$AB \sim \mathrm{MVN}[(I \times\cdot A)V_\beta;\ \Sigma \times\cdot A(X'X)^{-1}A']$$

$$ABC \sim \mathrm{MVN}[(C' \times\cdot I \times\cdot A)V_\beta;\ C'\Sigma C \times\cdot A(X'X)^{-1}A'].$$

These values are useful in constructing interval estimates around hypothesized values of transformations of β.

Sums of Squares Due to Regression

From Equation 2.6 we have

$$SS_E = Y'Y - B'X'Y$$

$$= Y'Y - B'X'XB.$$

The right-most terms in these expressions give the SSCP due to regression of Y on X. Typically, we let

$$SS_H = B'X'XB \qquad [2.10]$$

and call SS_H the matrix of SSCP explained by the hypothesis that the full model, given in Equation [2.1] holds for the data. SS_E

is that component of the total SSCP, $\underset{\sim}{Y}'\underset{\sim}{Y}$, that is left unexplained by the model.

Tests of Hypothesis Concerning $\underset{\sim}{\beta}$

Another way of defining $\underset{\sim\sim}{SS}_H$ is to say that it is the difference in $\underset{\sim\sim}{SS}_E$ obtained by the model given in Equation 2.1 and the model defined under the null hypothesis,

$$H_0: \underset{\sim}{\beta} = \underset{\sim}{0}.$$

Usually, however, we are not particularly interested in the hypothesis that all elements in $\underset{\sim}{\beta}$ are equal to zero; instead our interest is usually focused upon various subhypotheses involving linear transformations of the $\underset{\sim}{\beta}$. For instance, let the q × q matrix, $\underset{\sim}{A}$, be partitioned as

$$\underset{\sim}{A}' = [\underset{\sim}{A}_1' \mid \underset{\sim}{A}_2'],$$

where $\underset{\sim}{A}_1$ is (r × q) and $\underset{\sim}{A}_2$ is [(q − r) × q]. Then the model given in Equation 2.1 may be written as

$$\begin{aligned} \underset{\sim}{Y} &= \underset{\sim}{X}\underset{\sim}{A}\underset{\sim}{\beta} + \underset{\sim}{E} \\ &= \underset{\sim}{X}_1\underset{\sim}{A}_1\underset{\sim}{\beta} + \underset{\sim}{X}_2\underset{\sim}{A}_2\underset{\sim}{\beta} + \underset{\sim}{E} \qquad [2.11] \end{aligned}$$

where the partitioning of $\underset{\sim}{X}$ is in conformance with the partitioning of $\underset{\sim}{A}$. We now let

$$\underset{\sim}{A}_1\underset{\sim}{\beta} = \underset{\sim}{\beta}_1; \quad \underset{\sim}{A}_2\underset{\sim}{\beta} = \underset{\sim}{\beta}_2$$

so that if we wished, we could write the model as

$$\underset{\sim}{Y} = [\underset{\sim}{X}_1 \quad \underset{\sim}{X}_2] \begin{bmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \end{bmatrix} + E,$$

where we note that $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ do not necessarily represent direct partitions of the original $\boldsymbol{\beta}$. If we let

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix}$$

then LS estimates of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are given by

$$\mathbf{B}_1 = \mathbf{X}_{11}^{-1}\mathbf{X}_1'(\mathbf{Y} - \mathbf{X}_2\mathbf{B}_2),$$

$$\mathbf{B}_2 = \mathbf{X}_{22}^{-1}\mathbf{X}_2'(\mathbf{Y} - \mathbf{X}_1\mathbf{B}_1). \qquad [2.12]$$

Our interest lies in testing the subhypothesis that $\boldsymbol{\beta}_1 = \mathbf{A}_1\boldsymbol{\beta} = \mathbf{0}$, say. More formally, then, the hypothesis is

$$H_0: \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\beta}_2^* \end{bmatrix} \qquad [2.13]$$

where $\boldsymbol{\beta}_2$ is allowed to take on some (unrestricted) value under the hypothesis. Under H_0, the model is

$$\mathbf{Y} = \mathbf{X}_2\boldsymbol{\beta}_2^* + \mathbf{E}^*. \qquad [2.14]$$

The LS estimate of $\boldsymbol{\beta}_2^*$ is given by

$$\mathbf{B}_2^* = \mathbf{X}_{22}^{-1}\mathbf{X}_2'\mathbf{Y}. \qquad [2.15]$$

By steps similar to those given earlier we compute

$$\mathbf{SS}_{E*} = \mathbf{Y}'\mathbf{Y} - \mathbf{B}_2^{*\prime}\mathbf{X}_2'\mathbf{X}_2\mathbf{B}_2^*, \qquad [2.16]$$

where we let the last term on the right be denoted by $\mathbf{SS}_{H*}$, the SSCP explained by the restricted model under H_0 (viz, the

model in Equation 2.14).

Our interest is in the difference,

$$\underset{\sim\sim}{SS}_{H_0} = \underset{\sim\sim}{SS}_H - \underset{\sim\sim}{SS}_{H*} \qquad [2.17]$$

where $\underset{\sim\sim}{SS}_H$ was given in Equation 2.10. This difference is given by

$$\underset{\sim\sim}{SS}_{H_0} = \underset{\sim}{B}'\underset{\sim}{X}'\underset{\sim\sim}{XB} - \underset{\sim}{B}_2^{*\prime}\underset{\sim}{X}_2'\underset{\sim}{X}_2\underset{\sim}{B}_2^* \qquad [2.18]$$

$$= [(\underset{\sim}{B}_1 - \underset{\sim}{0})' \ (\underset{\sim}{B}_2 - \underset{\sim}{B}_2^*)'] \begin{bmatrix} \underset{\sim}{X}_{11} & \underset{\sim}{X}_{12} \\ \underset{\sim}{X}_{21} & \underset{\sim}{X}_{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{B}_1 - \underset{\sim}{0} \\ \underset{\sim}{B}_2 - \underset{\sim}{B}_2^* \end{bmatrix}$$

Before continuing, we note that from Equations 2.12 and 2.15

$$(\underset{\sim}{B}_2 - \underset{\sim}{B}_2^*) = \underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_2'\underset{\sim}{Y} - \underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_{21}\underset{\sim}{B}_1 - \underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_2'\underset{\sim}{Y}$$

$$= - \underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_{21}\underset{\sim}{B}_1.$$

Following through with the matrix operations in Equation 2.18 leads to the result

$$\underset{\sim\sim}{SS}_{H_0} = \underset{\sim}{B}_1'(\underset{\sim}{X}_{11}\underset{\sim}{B}_1 - \underset{\sim}{X}_{12}\underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_{21}\underset{\sim}{B}_1)$$

$$= \underset{\sim}{B}_1'(\underset{\sim}{X}_{11} - \underset{\sim}{X}_{12}\underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_{21})\underset{\sim}{B}_1.$$

Note, however, that from matrix theory we have the important result that the inverse of the term in the parentheses above is equivalent to $\underset{\sim}{A}_1(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{A}_1'$. That is to say

$$(\underset{\sim}{X}_{11} - \underset{\sim}{X}_{12}\underset{\sim}{X}_{22}^{-1}\underset{\sim}{X}_{21})^{-1} = \underset{\sim}{A}_1(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{A}_1'.$$

Therefore, we have

$$SS_{H_0} = B_1'[A_1(X'X)^{-1}A_1']^{-1}B_1$$

$$= B'A_1'[A_1(X'X)^{-1}A_1']^{-1}A_1B.$$

SS_{H_0}, as defined here, is the SSCP explained by the hypothesis

$$H_0: A_1\beta = 0. \qquad [2.20]$$

In many cases, we are interested in tests of hypotheses involving transformations on the columns of β. In other words, we are interested in hypotheses of the form

$$H_0: BC = 0. \qquad [2.21]$$

Under this hypothesis, the model in Equation [2.1] becomes

$$YC = X\beta C + EC$$

and the LS estimate of βC is given by

$$BC = (X'X)^{-1}X'YC$$

from which we can obtain

$$SS_{EC} = C'Y'YC - C'B'X'XBC$$

$$= C'SS_TC - C'SS_HC. \qquad [2.22]$$

where SS_H was defined in Equation 2.10

Similarly, it could be shown that for the MGLH,

$$H_0: A_1\beta C = 0$$

we have

$$SS_{EC} = C'SS_TC - C'B'A_1'[A_1(X'X)^{-1}A_1']^{-1}ABC$$

$$= C'SS_TC - C'SS_{H_0}C$$

$$= C'[SS_T - SS_{H_0}]C. \qquad [2.23]$$

We may summarize the results obtained thus far in this section by writing $\underset{\sim}{H}$ for $\underset{\sim}{C}'\underset{\sim}{SS}_{H_0}\underset{\sim}{C}$, $\underset{\sim}{E}$ for $\underset{\sim}{C}'\underset{\sim}{SS}_{EC}\underset{\sim}{C}$, and $\underset{\sim}{A}$ for $\underset{\sim}{A}_1$, and noting that the most general statement of an MGLH may be written as

$$\underset{\sim}{A}\underset{\sim}{\beta}\underset{\sim}{C} = \underset{\sim}{0}$$

where $\underset{\sim}{A}$ is a $g \times r$ matrix of rank g whose elements, a_{ij} ($i = 1, 2, \ldots, g$; $j = 1, 2, \ldots, r$) are used to select particular combinations from the rows of $\underset{\sim}{B}$; $\underset{\sim}{C}$ is a $p \times u$ matrix of rank u whose elements, c_{ij} ($i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, u$), are used to select linear combinations among the columns of $\underset{\sim}{\beta}$.

Test Criteria

Multivariate test criteria are usually a function of the characteristic roots of $\underset{\sim}{H}\underset{\sim}{E}^{-1}$ (or, equivalently, of the determinantal equation $|\underset{\sim}{H} - \lambda\underset{\sim}{E}| = 0$). Three popular test criteria are given in the next three paragraphs.

Trace criterion. The trace criterion is the trace of $\underset{\sim}{H}\underset{\sim}{E}^{-1}$ which is equivalent to the sum of the roots of the determinantal equation,

$$|\underset{\sim}{H} - \lambda_i\underset{\sim}{E}| = 0; \quad i = 1, 2, \ldots, u$$

$$\lambda_1 > \lambda_2 > \cdots > \lambda_u.$$

According to Anderson (1958; p. 224) the asymptotic distribution of $N \times (\text{trace } HE^{-1})$ is the chi square distribution with gu degrees of freedom (where g = rank of $\underset{\sim}{A}$, and u = rank of $\underset{\sim}{C}$).

Greatest characteristic root criterion. This test statistic uses the largest characteristic root, λ_1, of $\underset{\sim}{HE}^{-1}$. For convenience, λ_1 may be converted to

$$\theta = \frac{\lambda_1}{1 + \lambda_1}$$

for which tabled percentage points have been given (Heck [1960]; these have also been reproduced by Morrison [1967]). Parameters for entering the Heck tables are,

$$s = \min (g, u)$$

$$m = (|g - u| - 1)/2$$

$$n = (N - r - u - 1)/2, \text{ where } r = \text{rank of } \underset{\sim}{X}.$$

Wilks maximum likelihood criterion. This criterion makes use of the statistic

$$\Lambda = \left[\prod_{i=1}^{u} (1 + \lambda_i) \right]^{-1}$$

where λ_i $(i = 1, 2, \ldots, u)$ are, again, the characteristic roots of $\underset{\sim}{HE}^{-1}$. An equivalent form of the above expression is

$$\Lambda = \frac{|\underset{\sim}{E}|}{|\underset{\sim}{H} + \underset{\sim}{E}|}$$

With large N, a chi square test due to Bartlett (1951) is available. Thus,

$$\chi^2 = -[N - r - .5(u - g + 1)] \ln \Lambda$$

is distributed as chi square with gu degrees of freedom. A better approximation (Rao [1965]) is given by

$$F = \frac{1 - \Lambda^{1/s}_{u,g,N-r}}{\Lambda^{1/s}_{u,g,N-r}} \cdot \frac{S[(N-r) + g-u-1] - .5(gu-2)}{gu}$$

where

$$S = \sqrt{\frac{g^2u^2 - 4}{g^2 + u^2 - 5}}$$

Under the hypothesis, F is approximately distributed as an F statistic with gu and {S [(N-r) + g-u-1] -.5(gu-2)} degrees of freedom.

For certain values of g and u, Anderson (1958, §8.5.1) gives the appropriate exact F statistic. For example when g = 1,

$$F = \frac{1 - \Lambda_{u,1,N-r}}{\Lambda_{u,1,N-r}} \cdot \frac{N-r+1-u}{u} = F_{u,N-r+1-u}.$$

In the special case where the rank of A (or the rank of $\underset{\sim}{C}$) is equal to one, the product $\underset{\sim\sim}{HE}^{-1}$, has only one non-zero characteristic root. In this case, the largest root, the sum of the roots, and the product of the (non-zero) roots are all the same, thus, making the three criteria equivalent.

3. *Typical Applications*

Having described the basic equations of the MGLH, we now turn to an algebraic exposition of some of the more typical models found in educational research. Examples of typical applications using real data can be found in many of the references cited earlier as well as in Olson (1971). In the discussion which follows, we begin with the model for the analysis of variance, move to a brief discussion of the analysis of covariance, and finally present a discussion on the analysis of repeated measures designs.

Multivariate Analysis of Variance

Let the configuration of Figure 1 represent the general design of a 2 × 3 factorial experiment. There are n_{ij} experimental units in each cell, and for purposes of exposition it will be assumed that $n_{11} = n_{12} = \ldots = n_{23} = n$, though we realize this is *not* a necessary restriction on the general linear model. Measurements on each of p dependent measures ($Y^{(\ell)}$; $\ell = 1, 2, \ldots, p$) have been collected on each of the 6 × n experimental units. The notation, $Y_{kij}^{(\ell)}$, denotes the measurement of the ℓ'th dependent variable for the k'th subject in the ij'th treatment combination. It is assumed that each of the $N = 6 \times n$ vectors, $\underset{\sim}{Y}_k^{(\ell)}$ ($\ell = 1, 2, \ldots, p$), have been independently sampled and follow the multinormal law with expected values, $\mathcal{E}(\underset{\sim}{Y}_k)$, and common variance-covariance matrix $\underset{\sim}{\Sigma}$. It is possible to write the N × p observations, given in Figure 1 as the N × p supermatrix of observations,

$$\underset{\sim}{Y} = \begin{bmatrix} {}_n\underset{\sim}{Y}_{11}^{(1)} & {}_n\underset{\sim}{Y}_{11}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{11}^{(p)} \\ {}_n\underset{\sim}{Y}_{12}^{(1)} & {}_n\underset{\sim}{Y}_{12}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{12}^{(p)} \\ {}_n\underset{\sim}{Y}_{13}^{(1)} & {}_n\underset{\sim}{Y}_{13}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{13}^{(p)} \\ {}_n\underset{\sim}{Y}_{21}^{(1)} & {}_n\underset{\sim}{Y}_{21}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{21}^{(p)} \\ {}_n\underset{\sim}{Y}_{22}^{(1)} & {}_n\underset{\sim}{Y}_{22}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{22}^{(p)} \\ {}_n\underset{\sim}{Y}_{23}^{(1)} & {}_n\underset{\sim}{Y}_{23}^{(2)} & \cdots & {}_n\underset{\sim}{Y}_{23}^{(p)} \end{bmatrix}$$

$Y_{111}^{(1)}\ Y_{111}^{(2)}\ \cdots\ Y_{111}^{(p)}$	$Y_{112}^{(1)}\ Y_{112}^{(2)}\ \cdots\ Y_{112}^{(p)}$	$Y_{113}^{(1)}\ Y_{113}^{(2)}\ \cdots\ Y_{113}^{(p)}$
$Y_{211}^{(1)}\ Y_{211}^{(2)}\ \cdots\ Y_{211}^{(p)}$	$Y_{212}^{(1)}\ Y_{212}^{(2)}\ \cdots\ Y_{212}^{(p)}$	$Y_{213}^{(1)}\ Y_{213}^{(2)}\ \cdots\ Y_{213}^{(p)}$
... ...	... ...	... ...
$Y_{n11}^{(1)}\ Y_{n11}^{(2)}\ \cdots\ Y_{n11}^{(p)}$	$Y_{n12}^{(1)}\ Y_{n12}^{(2)}\ \cdots\ Y_{n12}^{(p)}$	$Y_{n13}^{(1)}\ Y_{n13}^{(2)}\ \cdots\ Y_{n13}^{(p)}$
$Y_{121}^{(1)}\ Y_{121}^{(2)}\ \cdots\ Y_{121}^{(p)}$	$Y_{122}^{(1)}\ Y_{122}^{(2)}\ \cdots\ Y_{122}^{(p)}$	$Y_{123}^{(1)}\ Y_{123}^{(2)}\ \cdots\ Y_{123}^{(p)}$
$Y_{221}^{(1)}\ Y_{221}^{(2)}\ \cdots\ Y_{221}^{(p)}$	$Y_{222}^{(1)}\ Y_{222}^{(2)}\ \cdots\ Y_{222}^{(p)}$	$Y_{223}^{(1)}\ Y_{223}^{(2)}\ \cdots\ Y_{223}^{(p)}$
... ...	... ...	... ...
$Y_{n21}^{(1)}\ Y_{n21}^{(2)}\ \cdots\ Y_{n21}^{(p)}$	$Y_{n22}^{(1)}\ Y_{n22}^{(2)}\ \cdots\ Y_{n22}^{(p)}$	$Y_{n23}^{(1)}\ Y_{n23}^{(2)}\ \cdots\ Y_{n23}^{(p)}$

Figure 1: Diagram of a General 2×3 Factorial Experiment.

where n denotes the number of rows in each of the n × 1 submatrices. An appropriate design matrix, describing the experimental design, would be the N × 6 Helmert-type matrix, of rank 6,

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} \\ \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{1} \\ \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{0} & \underset{\sim}{n}^{-2} & \underset{\sim}{n}^{0} & \underset{\sim}{n}^{-2} \\ \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{-1} \\ \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} \\ \underset{\sim}{n}^{1} & \underset{\sim}{n}^{-1} & \underset{\sim}{n}^{0} & \underset{\sim}{n}^{-2} & \underset{\sim}{n}^{0} & \underset{\sim}{n}^{2} \end{bmatrix}.$$

It can easily be shown that the solution of the normal equations (Equation 2.4) yields the matrix of estimates, $\underset{\sim}{B}$, shown in Figure 2. It is obvious that the first row in $\underset{\sim}{B}$ estimates the grand means of the dependent variates. The second row estimates the row effect. The third and fourth rows (considered jointly) estimate the column effect, and the final two rows, the row × column interaction[1].

In multivariate analysis of variance, hypotheses of interest commonly involve only effects due to independent variables (i.e., ways of classification). Hypotheses concerning contrasts among dependent variables are considered under profile analysis. When all dependent variables are to be included, the matrix $\underset{\sim}{C}$, of Equation 2.24 is set equal to $\underset{\sim}{I}$, the identity matrix. To test hypotheses of no treatment or interaction effects, the following A-matrices are constructed:

---

[1]The reader should be aware that when cell frequencies are disproportional (i.e., the design is unbalanced) estimates of row, column, and interaction effects will be confounded. See Overall & Spiegl (1969) or Searle (1971; pp. 138-139) for procedures on treating unbalanced designs.

Film as One Page

$$\underset{\sim}{B} = \begin{bmatrix} (\bar{Y}^{(1)}_{\cdot\cdot\cdot}) & (\bar{Y}^{(2)}_{\cdot\cdot\cdot}) & \cdots & (\bar{Y}^{(P)}_{\cdot\cdot\cdot}) \\ \dfrac{(\bar{Y}^{(1)}_{\cdot 1 \cdot}-\bar{Y}^{(1)}_{\cdot 2 \cdot})}{2} & \dfrac{(\bar{Y}^{(2)}_{\cdot 1 \cdot}-\bar{Y}^{(2)}_{\cdot 2 \cdot})}{2} & \cdots & \dfrac{(\bar{Y}^{(P)}_{\cdot 1 \cdot}-\bar{Y}^{(P)}_{\cdot 2 \cdot})}{2} \\ \dfrac{(\bar{Y}^{(1)}_{\cdot\cdot 1}-\bar{Y}^{(1)}_{\cdot\cdot 2})}{2} & \dfrac{(\bar{Y}^{(2)}_{\cdot\cdot 1}-\bar{Y}^{(2)}_{\cdot\cdot 2})}{2} & \cdots & \dfrac{(\bar{Y}^{(P)}_{\cdot\cdot 1}-\bar{Y}^{(P)}_{\cdot\cdot 2})}{2} \\ \dfrac{(\bar{Y}^{(1)}_{\cdot\cdot 1}+\bar{Y}^{(1)}_{\cdot\cdot 2}-2\bar{Y}^{(1)}_{\cdot\cdot 3})}{6} & \dfrac{(\bar{Y}^{(2)}_{\cdot\cdot 1}+\bar{Y}^{(2)}_{\cdot\cdot 2}-2\bar{Y}^{(2)}_{\cdot\cdot 3})}{6} & \cdots & \dfrac{(\bar{Y}^{(P)}_{\cdot\cdot 1}+\bar{Y}^{(P)}_{\cdot\cdot 2}-2\bar{Y}^{(P)}_{\cdot\cdot 3})}{6} \\ \dfrac{(\bar{Y}^{(1)}_{\cdot 11}-\bar{Y}^{(1)}_{\cdot 12}-\bar{Y}^{(1)}_{\cdot 21}+\bar{Y}^{(1)}_{\cdot 22})}{4} & \dfrac{(\bar{Y}^{(2)}_{\cdot 11}-\bar{Y}^{(2)}_{\cdot 12}-\bar{Y}^{(2)}_{\cdot 21}+\bar{Y}^{(2)}_{\cdot 22})}{4} & \cdots & \dfrac{(\bar{Y}^{(P)}_{\cdot 11}-\bar{Y}^{(P)}_{\cdot 12}-\bar{Y}^{(P)}_{\cdot 21}+\bar{Y}^{(P)}_{\cdot 22})}{4} \\ \dfrac{(\bar{Y}^{(1)}_{\cdot 11}+\bar{Y}^{(1)}_{\cdot 12}-2\bar{Y}^{(1)}_{\cdot 13}-\bar{Y}^{(1)}_{\cdot 21}-\bar{Y}^{(1)}_{\cdot 22}+2\bar{Y}^{(1)}_{\cdot 23})}{12} & \dfrac{(\bar{Y}^{(2)}_{\cdot 11}+\bar{Y}^{(2)}_{\cdot 12}-2\bar{Y}^{(2)}_{\cdot 13}-\bar{Y}^{(2)}_{\cdot 21}-\bar{Y}^{(2)}_{\cdot 22}+2\bar{Y}^{(2)}_{\cdot 23})}{12} & \cdots & \dfrac{(\bar{Y}^{(P)}_{\cdot 11}+\bar{Y}^{(P)}_{\cdot 12}-2\bar{Y}^{(P)}_{\cdot 13}-\bar{Y}^{(P)}_{\cdot 21}-\bar{Y}^{(P)}_{\cdot 22}+2\bar{Y}^{(P)}_{\cdot 23})}{12} \end{bmatrix}$$

Figure 2. Solution o the normal equations for $\underset{\sim}{B}$.

$$\underset{\sim}{A}_{(M)} = [1 \; 0 \; 0 \; 0 \; 0 \; 0]$$

$$\underset{\sim}{A}_{(R)} = [0 \; 1 \; 0 \; 0 \; 0 \; 0]$$

$$\underset{\sim}{A}_{(C)} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \qquad [3.1]$$

$$\underset{\sim}{A}_{(RC)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

By substituting these matrices into the general linear hypothesis framework and letting $\underset{\sim}{C} = \underset{\sim}{I}$, the following hypotheses are then tested:

$$H_{0,M}: \; \underset{\sim}{A}_{(M)} \underset{\sim}{\beta} \underset{\sim}{C} = \underset{\sim}{0}$$

$$H_{0,R}: \; \underset{\sim}{A}_{(R)} \underset{\sim}{\beta} \underset{\sim}{C} = \underset{\sim}{0} \qquad [3.2]$$

$$H_{0,C}: \; \underset{\sim}{A}_{(C)} \underset{\sim}{\beta} \underset{\sim}{C} = \underset{\sim}{0}$$

$$H_{0,AC}: \; \underset{\sim}{A}_{(RC)} \underset{\sim}{\beta} \underset{\sim}{C} = \underset{\sim}{0}$$

where

$H_{0,M}$ is the null hypothesis that the overall grand means equal zero;

$H_{0,R}$ is the null hypothesis of no row effect;

$H_{0,C}$ is the null hypothesis of no column effect; and

$H_{0,RC}$ is the null hypothesis of no interaction.

Ar alternative way to have constructed the design matrix is

$$X = \begin{bmatrix} \underset{\sim}{n}^1 & & & & & \\ & \underset{\sim}{n}^1 & & & \underset{\sim}{0} & \\ & & \underset{\sim}{n}^1 & & & \\ & & & \underset{\sim}{n}^1 & & \\ & \underset{\sim}{0} & & & \underset{\sim}{n}^1 & \\ & & & & & \underset{\sim}{n}^1 \end{bmatrix}, \qquad [3.3]$$

where again n denotes the number of rows in each of the $n \times 1$ submatrices. Using this design matrix, the solution of the normal equations would yield

$$\underset{\sim}{B} = \begin{bmatrix} \bar{Y}_{11}^{(1)} & \bar{Y}_{11}^{(2)} & \cdots & \bar{Y}_{11}^{(p)} \\ \bar{Y}_{12}^{(1)} & \bar{Y}_{12}^{(2)} & \cdots & \bar{Y}_{12}^{(p)} \\ \bar{Y}_{13}^{(1)} & \bar{Y}_{13}^{(2)} & \cdots & \bar{Y}_{13}^{(p)} \\ \bar{Y}_{21}^{(1)} & \bar{Y}_{21}^{(2)} & \cdots & \bar{Y}_{21}^{(p)} \\ \bar{Y}_{22}^{(1)} & \bar{Y}_{22}^{(2)} & \cdots & \bar{Y}_{22}^{(p)} \\ \bar{Y}_{23}^{(1)} & \bar{Y}_{23}^{(2)} & \cdots & \bar{Y}_{23}^{(p)} \end{bmatrix} \qquad [3.4]$$

the group means on each dependent variable. Using $\underset{\sim}{B}$ as defined above the A matrices,

$$\begin{aligned} \underset{\sim}{A}_{(M)} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \\ \underset{\sim}{A}_{(R)} &= \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix} \\ \underset{\sim}{A}_{(C)} &= \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 0 \\ 1 & 1 & -2 & 1 & 1 & -2 \end{bmatrix} \\ \underset{\sim}{A}_{(RC)} &= \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 \\ 1 & 1 & -2 & -1 & -1 & 2 \end{bmatrix} \end{aligned} \qquad [3.5]$$

may then be used to test the hypotheses given in Equation 3.2. Significance tests are computed by first finding $\underset{\sim}{H}$ and $\underset{\sim}{E}$ and then applying one of the test criteria given earlier.

Analysis of Covariance

Suppose that, prior to experimental treatment, measurements on the covariates, $Z^{(1)}$ and $Z^{(2)}$, were collected for each experimental unit in Figure 1. An appropriate matrix of independent variables would then be constructed of both design variables and the predictor variables, $Z^{(1)}$, $Z^{(2)}$. Thus, corresponding to [3.3],

$$\underset{\sim}{X} = \left[\begin{array}{cccccc|cc} \underset{\sim}{n}^1 & & & & & & \underset{\sim}{Z}_{11}^{(1)} & \underset{\sim}{Z}_{11}^{(2)} \\ & \underset{\sim}{n}^1 & & \underset{\sim}{0} & & & \underset{\sim}{Z}_{12}^{(1)} & \underset{\sim}{Z}_{12}^{(2)} \\ & & \underset{\sim}{n}^1 & & & & \underset{\sim}{Z}_{13}^{(1)} & \underset{\sim}{Z}_{13}^{(2)} \\ & & & \underset{\sim}{n}^1 & & & \underset{\sim}{Z}_{21}^{(1)} & \underset{\sim}{Z}_{21}^{(2)} \\ & & & & \underset{\sim}{n}^1 & & \underset{\sim}{Z}_{22}^{(1)} & \underset{\sim}{Z}_{22}^{(2)} \\ \underset{\sim}{0} & & & & & \underset{\sim}{n}^1 & \underset{\sim}{Z}_{23}^{(1)} & \underset{\sim}{Z}_{23}^{(2)} \end{array}\right] \qquad [3.6]$$

where the notation, $\underset{\sim}{Z}_{ij}^{(\ell)}$, denotes the $n \times 1$ vector of measurements, within the ij'th level of classification, on the ℓ'th covariate. The solution of the normal equations yields the $8 \times p$ matrix of adjusted means and regression coefficients

$$\underset{\sim}{B}^* = \left[\begin{array}{cccc} \bar{Y}_{11}^{*(1)} & \bar{Y}_{11}^{*(2)} & \cdots & \bar{Y}_{11}^{*(p)} \\ \bar{Y}_{12}^{*(1)} & \bar{Y}_{12}^{*(2)} & \cdots & \bar{Y}_{12}^{*(p)} \\ \cdots & \cdots & \cdots & \cdots \\ \bar{Y}_{23}^{*(1)} & \bar{Y}_{23}^{*(2)} & \cdots & \bar{Y}_{23}^{*(p)} \\ \hline w_1^{(1)} & w_1^{(2)} & \cdots & w_1^{(p)} \\ w_2^{(1)} & w_2^{(2)} & \cdots & w_2^{(p)} \end{array}\right] \qquad [3.7]$$

where $w_\ell^{(1)}, w_\ell^{(2)}, \ldots, w_\ell^{(p)}$ denote the within-class regression coefficients of $Y^{(1)}, Y^{(2)}, \ldots, Y^{(p)}$ on $Z^{(\ell)}$. By constructing the A-matrices,

$$\underset{\sim}{A}_{(M)} = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1 \;\vdots\; 0 \quad 0]$$

$$\underset{\sim}{A}_{(R)} = [1 \quad 1 \quad 1 \quad -1 \quad -1 \quad -1 \;\vdots\; 0 \quad 0]$$

$$\underset{\sim}{A}_{(C)} = \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 0 & \vdots & 0 & 0 \\ 1 & -1 & -2 & 1 & 1 & -2 & \vdots & 0 & 0 \end{bmatrix} \qquad [3.8]$$

$$\underset{\sim}{A}_{(RC)} \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 & \vdots & 0 & 0 \\ 1 & 1 & -2 & -1 & -1 & 2 & \vdots & 0 & 0 \end{bmatrix},$$

and letting $\underset{\sim}{C} = \underset{\sim}{I}$, the hypotheses (Equation 3.2) can then be tested, this time for effects adjusted for regression on $Z^{(1)}$ and $Z^{(2)}$.

If the subgroups on the covariates are computed beforehand, then tests of unadjusted effects can be obtained easily. We note that any element in $\underset{\sim}{B}^*$, $B_{ij}^{*(\ell)}$ say, is the coefficient for the regression of $Y^{(\ell)}$ on X_i adjusted for regression on the $Z^{(\ell)}$. For instance, the cell mean for group 1,2 on the first dependent variable is given by $\bar{Y}_{12}^{*(1)} = \bar{Y}_{12}^{(1)} - W_1\bar{Z}^{(1)} - W_2\bar{Z}^{(2)}$. It follows that the *unadjusted* cell mean for this group is

$$\bar{Y}_{12}^{(1)} = \bar{Y}_{12}^{*(1)} + W_1^{(1)}\bar{Z}^{(1)} + W_2^{(1)}\bar{Z}^{(2)}.$$

Therefore the A-matrix,

$$\underset{\sim}{A} = [0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \;\vdots\; \bar{Z}^{(1)} \quad \bar{Z}^{(2)}]$$

would yield

$$\bar{Y}^{*(\ell)}_{12} + [\bar{Z}^{(1)}W_1^{(1)} + \bar{Z}^{(2)}W_2^{(1)}]$$

$$= \bar{Y}^{(\ell)}_{12} - [W_1^{(1)}\bar{Z}^{(1)} + W_2^{(1)}Z^{(2)}] + [W_1^{(1)}Z^{(1)} + W_2^{(1)}Z^{(2)}]$$

$$= Y^{(\ell)}_{12} \quad (\text{for } \ell = 1, 2, \ldots, p),$$

the unadjusted cell means. By including the overall covariate means in the last two columns of each of the A-matrices given earlier, unadjusted treatment effects can be easily computed.

Profile Analysis (Repeated Measures)

Profile analysis has been aptly discussed by Marks (1968; see also Morrison [1967; pp. 168-197]). Essentially the problem is one of determining whether the shape of a mean vector is equal to that of another mean vector. In the two-way classification, for example, the problem is to determine whether mean vectors are equal for the different ways of classification. This is equivalent to asking whether contrasts among selected dependent measures for one group are equal to the same contrasts for another group. Observant readers will notice that this is precisely the problem in univariate repeated measures analysis of variance where the measures on the multiple dependent variables are taken over time. In the example situation described in Figure 1, the researcher may wish to determine whether differences among adjacent dependent variables can be considered equal for all levels of the column factor. That is, he may wish to simultaneously test,

$$\begin{bmatrix} \mu^{(1)}_{..1} - \mu^{(2)}_{..1} \\ \mu^{(2)}_{..1} - \mu^{(3)}_{..1} \\ \cdots \quad \cdots \\ \mu^{(p-1)}_{..1} - \mu^{(p)}_{..1} \end{bmatrix} = \begin{bmatrix} \mu^{(1)}_{..2} - \mu^{(2)}_{..2} \\ \mu^{(2)}_{..2} - \mu^{(3)}_{..2} \\ \cdots \quad \cdots \\ \mu^{(p-1)}_{..2} - \mu^{(p)}_{..2} \end{bmatrix}$$

and

$$\begin{bmatrix} (\mu^{(1)}_{..1} + \mu^{(1)}_{..2}) - (\mu^{(2)}_{..1} + \mu^{(2)}_{..2}) \\ (\mu^{(2)}_{..1} + \mu^{(2)}_{..2}) - (\mu^{(3)}_{..1} + \mu^{(3)}_{..2}) \\ \cdots \\ (\mu^{(p-1)}_{...1} + \mu^{(p-1)}_{...2}) - (\mu^{(p)}_{..1} + \mu^{(p)}_{..2}) \end{bmatrix} = 2 \times \begin{bmatrix} \mu^{(1)}_{..3} - \mu^{(2)}_{..3} \\ \mu^{(2)}_{..3} - \mu^{(3)}_{..3} \\ \cdots \\ \mu^{(p-1)}_{..3} - \mu^{(p)}_{..3} \end{bmatrix},$$

where the $\mu^{(\ell)}_{..j}$'s represent population values. It can be easily shown that with either of the design matrices given earlier, its corresponding solution to the normal equations, the appropriate set of A-matrices, and the p × (p-1) matrix,

$$\underset{\sim}{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & \cdots & 0 & -1 \end{bmatrix};$$

then the MGLH

$$\underset{\sim}{A}_c \underset{\sim}{\beta} \underset{\sim}{C} = \underset{\sim}{0}$$

will yield the appropriate hypotheses for tests on profiles or repeated measures.

Repeated Measures with Covariates

Two distinct types of repeated measures design involving covariates can be identified (cf., Winer [1962, 1971]). In the first, measures on the covariates are collected prior to the onset of any treatment. In the second, covariate measures are collected concomitantly, in time, with measures on the dependent variable. Of course, a mixture of the two types of designs is possible. We shall consider each type of design separately.

Covariates prior to the onset of treatment. The first type of covariate repeated measures design might be diagrammed as follows,

$$\begin{array}{cccc|cccc}
\multicolumn{8}{c}{\text{- - - - - - - - - - - - - Time - - - - - - - - - - - - - -}\rightarrow} \\
\underset{\sim}{Z}_1^{(1)} & \underset{\sim}{Z}_1^{(2)} & \cdots & \underset{\sim}{Z}_1^{(h)} & \underset{\sim}{Y}_1^{(1)} & \underset{\sim}{Y}_1^{(2)} & \cdots & \underset{\sim}{Y}_1^{(p)} \\
\underset{\sim}{Z}_2^{(1)} & \underset{\sim}{Z}_2^{(2)} & \cdots & \underset{\sim}{Z}_2^{(h)} & \underset{\sim}{Y}_2^{(1)} & \underset{\sim}{Y}_2^{(2)} & \cdots & \underset{\sim}{Y}_2^{(p)}
\end{array}$$

where the vertical line indicates the beginning of treatment; the $\underset{\sim}{Z}_j^{(\ell)}$, covariate observations for group j; and the $Y_j^{(\ell)}$, measures on the dependent variables for group j. All vectors in the diagram are $n_j \times 1$, n_j being the number of subjects in group j. A good design matrix for this design would be the $(n_1 + n_2) \times (h + 2)$ matrix

$$\underset{\sim}{X} = \begin{bmatrix} {}_{n_j}\underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{Z}_1 \\ \underset{\sim}{0} & {}_{n_j}\underset{\sim}{1} & \underset{\sim}{Z}_2 \end{bmatrix} .$$

The $(n_1 + n_2) \times p$ matrix of dependent measures, $\underset{\sim}{Y}$, would be laid out in the usual way. The solution to the normal equations would yield the $(h + 2) \times p$ matrix of estimates,

$$
\underset{\sim}{B} = \begin{bmatrix}
\bar{Y}^{*(1)}_1 & \bar{Y}^{*(2)}_1 & \cdots & \bar{Y}^{*(p)}_1 \\
\bar{Y}^{*(1)}_2 & \bar{Y}^{*(2)}_2 & \cdots & \bar{Y}^{*(p)}_2 \\
W^{(1)}_1 & W^{(2)}_1 & \cdots & W^{(p)}_1 \\
\cdots & \cdots & \cdots & \cdots \\
W^{(1)}_h & W^{(2)}_h & \cdots & W^{(p)}_h
\end{bmatrix}
$$

where $\bar{Y}^{*(\ell)}_j$ is the j'th group mean on variable ℓ adjusted for all independent variables. This design presents no problem, and MGLHs can be constructed and tested in the usual manner.

Repeated measures with covariates measured concomitantly. The second type of repeated measures designs involving covariates is somewhat more difficult to handle. The design, for a two-group classification, can be diagrammed as follows,

$$
\begin{array}{cccccccc}
\multicolumn{2}{c}{\textit{Time 1}} & \multicolumn{2}{c}{\textit{Time 2}} & & \multicolumn{2}{c}{\textit{Time p}} \\
\underset{\sim}{Z}^{(1)}_1 & \underset{\sim}{Y}^{(1)}_1 & \underset{\sim}{Z}^{(2)}_1 & \underset{\sim}{Y}^{(2)}_1 & \cdots & \underset{\sim}{Z}^{(p)}_1 & \underset{\sim}{Y}^{(p)}_1 \\
\underset{\sim}{Z}^{(1)}_2 & \underset{\sim}{Y}^{(1)}_2 & \underset{\sim}{Z}^{(2)}_2 & \underset{\sim}{Y}^{(2)}_2 & \cdots & \underset{\sim}{Z}^{(p)}_2 & \underset{\sim}{Y}^{(p)}_2 ,
\end{array}
$$

where the $\underset{\sim}{Z}^{(\ell)}_j$ are the j'th-group observations on the covariate at time p; the $Y^{(\ell)}_j$ are the j'th-group measures on the dependent variable at time p. In this situation, the design matrix, as constructed in the previous situation would be inappropriate since it would lead to estimates of group means on each dependent variable adjusted for all independent variables *including* those

covariate measures which follow the dependent measures in time. What is desired, however, are estimates of group means, on each dependent variable, which are adjusted only for those covariate measures which were collected at the same point in time or earlier. In other words, the parameter matrix of interest is

$$\beta = \begin{bmatrix} \mu^{(1)}_{1\cdot 1} & \mu^{(2)}_{1\cdot 12} & \mu^{(3)}_{1\cdot 123} & \cdots & \mu^{(p)}_{1\cdot 12\ldots p} \\ \mu^{(1)}_{2\cdot 1} & \mu^{(2)}_{2\cdot 12} & \mu^{(3)}_{2\cdot 123} & \cdots & \mu^{(p)}_{2\cdot 22\ldots p} \\ \omega^{(1)}_{1} & \omega^{(2)}_{1\cdot 2} & \omega^{(3)}_{1\cdot 23} & \cdots & \omega^{(p)}_{1\cdot 23\ldots p} \\ 0 & \omega^{(2)}_{2\cdot 1} & \omega^{(3)}_{2\cdot 13} & \cdots & \omega^{(p)}_{2\cdot 13\ldots p} \\ 0 & 0 & \omega^{(3)}_{3\cdot 12} & \cdots & \omega^{(p)}_{3\cdot 14\ldots p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \omega^{(p)}_{p\cdot 12\ldots p\,1} \end{bmatrix},$$

where

$\mu^{(\ell)}_{i\cdot jkl\ldots}$ = population group mean on dependent variable ℓ adjusted for $Z^{(i)}$, $Z^{(k)}$, $Z^{(l)}$, ...

$\omega^{(\ell)}_{i\cdot jk\ldots}$ = within-class population regression coefficient of $Y^{(\ell)}$ on $Z^{(i)}$ adjusted for $Z^{(j)}$, $Z^{(k)}$, ...

To obtain estimates of this matrix a *generalized* (see Press [1972; pp. 217-227]) multivariate linear model is used. This model has the form given in Figure 3.

$$
\begin{bmatrix}
Y_1^{(1)} & 0 & \cdots & 0 \\
0 & Y_1^{(2)} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & Y_1^{(p)} \\
Y_2^{(1)} & 0 & \cdots & 0 \\
0 & Y_2^{(2)} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & Y_2^{(p)}
\end{bmatrix}
=
\begin{bmatrix}
1 & 0 & Z_1^{(1)} & 0 & \cdots & 0 \\
1 & 0 & Z_1^{(1)} & Z_1^{(2)} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
1 & 0 & Z_1^{(1)} & Z_1^{(2)} & \cdots & Z_1^{(p)} \\
0 & 1 & Z_2^{(1)} & 0 & \cdots & 0 \\
0 & 1 & Z_2^{(1)} & Z_2^{(2)} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 1 & Z_2^{(1)} & Z_2^{(2)} & \cdots & Z_2^{(p)}
\end{bmatrix}
\beta + E
$$

Figure 3. Generalized Multivariate Linear Model

where the $Y_j^{(\ell)}$, $Z_j^{(\ell)}$, are $n_j \times 1$ vectors defined in the diagram for the design, and the symbols, 1 and 0, denote conformable column vectors of 1's and 0's respectively. In the model, the matrix, β, is the desired matrix of parameters.

The least squares solution to the model in Figure 3 provides the appropriate matrix of estimates. Tests of the repeated measures effects are then made by forming MGLHs involving the first two rows of B. For instance, with

$$A = [1 - 1 \;\; 0 \;\; 0 \;\; \ldots \;\; 0] \quad (\text{where } A \text{ is } 1 \times (p + 2))$$

$$C' = \begin{bmatrix}
1 & -1 & 0 & \cdots & 0 & 0 \\
0 & 1 & -1 & \cdots & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & 1 & -1
\end{bmatrix}, \quad (\text{where } C \text{ is } (p - 1) \times p)$$

the MGLH, $ABC = 0$, would provide an appropriate test of the null hypothesis of equal profiles for the two groups.

List of References

Anderson, T. W. An introduction to multivariate statistical analysis. New York: Wiley, 1958.

Bock, R. D. Multivariate analysis of variance of repeated measurements. In C. W. Harris (Ed.), Problems in Measuring Change. Madison: University of Wisconsin, 1963a.

Bock, R. D. Programming univariate and multivariate analysis of variance. Technometrics, 1963, 5, 95-117b.

Bock, R. D., & Haggard, E. A. The use of multivariate analysis of variance in behavioral research. In D. K. Whitla (Ed.), Handbook of Measurement and Assessment in Behavioral Sciences.

Cooley, W. W., & Lohnes, P. R. Multivariate Data Analysis. New York: Wiley, 1971.

Cornish, E. A. An application of the Kronecker product of matrices in multiple regression. Biometrics, 1951, 7, 1-116.

Draper, N. R., & Smith, H. Applied Regression Analysis. New York: Wiley, 1966.

Finn, J. D. A General Model for Multivariate Analysis. New York: Holt, 1974.

Harris, R. J. A Primer of Multivariate Statistics. New York: Academic Press, 1975.

Hummel, T. J. & Sligo, I. R. Empirical comparison of univariate and multivariate analysis of variance procedures. Psychological Bulletin, 1971, 76, 49-57.

Jones, L. V. Analysis of variance in its multivariate developments. In R. B. Cattell (Ed.), Handbook of Multivariate Experimental Psychology. Chicago: Rand-McNally, 1966.

Kempthorne, O. The Design and Analysis of Experiments. Huntington, New York: Robert E. Krieger, 1952.

Kulback, S. Information Theory and Statistics. New York: Dover, 1968.

Marks, E. Profile analysis in a two-way classification problem. Multivariate Behavioral Research, 1968, 3, 96-106.

Mendenhall, W. Introduction to Linear Models and the Design and Analysis of Experiments. Belmont, California: Wadsworth, 1968.

Morrison, D. F. Multivariate statistical methods. New York: McGraw-Hill, 1967.

Olson, G. H. A multivariate examination of the effects of behavioral objectives, knowledge of results, and the assignment of grades on the facilitation of classroom learning. Unpublished Doctoral Dissertation, Tallahassee: The Florida State University, 1971.

Press, S. J. Applied Multivariate Analysis. New York: Holt, 1972.

Rao, C. R. Linear Statistical Inference and its Applications. New York: Wiley, 1965.

Seal, H. Multivariate Statistical Analysis for Biologists. New York: Wiley, 1964.

Searle, S. R. Matrix Algebra for the Biological Sciences. New York: Wiley, 1966.

Seber, G. A. The Linear Hypothesis, Griffen's Statistical Monographs. New York: Hafner, 1966.

Smith, H., Gnanadesikan, R., & Hughes, J. B. Multivariate analysis of variance (MANOVA). Biometrics, 1962, 22-41.

Snow, R. E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 256-292.

Tatsuoka, M. M. Multivariate analysis: Techniques for educational and psychological research. New York: Wiley, 1971.

Tatsuoka, M. M. Multivariate Analysis in Educational Research. In F. N. Kerlinger (Ed.), Review of Research in Education, No. 1. Itasca, Ill.: Peacock (1973).

Van de Greer, J. P. Introduction to Multivariate Analysis for the Social Sciences. San Francisco: W. H. Freeman, 1971.

Vartak, M. H. On an application of Kronecker product of matrices to statistical designs. Annals of Mathematical Statistics, 1955, 26, 420-437.